Measuring and Modeling Human Probabilistic Segmentation Maps

Jonathan Vacher (jonathan.vacher@einsteinmed.org), Pascal Mamassian, Ruben Coen-Cagli

Einstein — Albert Einstein College of Medicine · Montefiore · cnrs · ENS · PSL

1. Introduction

Visual segmentation is a core function of biological vision:

- involves Gestalt principles, *e.g.* grouping by similarity, proximity and good continuation [1]
- visual cortical neurons are sensitive to those cues [2]

Purpose: to compare the prediction of different models to human performance

Feedforward models: comparing the local summary statistics of low-level visual features [3, 4].
Alternative view: perceptual segmentation emerges from probabilistic inference.

To do that:

- we propose a new protocol to measure segmentation maps and variability
- we measured segmentation maps of composite artificial images

2. Measure

A new task to measure segmentation maps:

- Ask the participant to decompose the image in K segments
- Show the image for 3 s
- Run a sequence of M trials: does the pair belong to the same segment ?

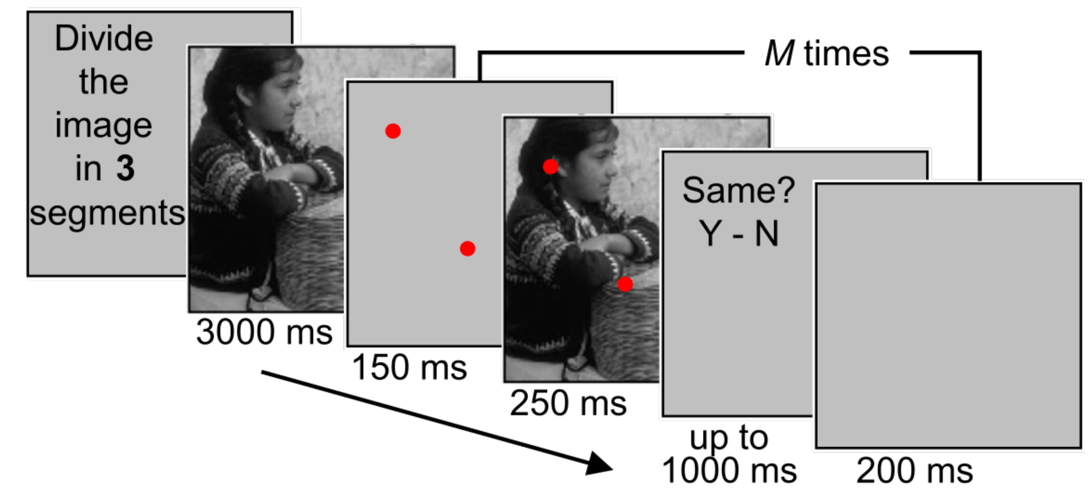


Figure 1: Experiment trial layout

3. Reconstruction

For any response model $p_{\mathbf{i},\mathbf{j}}(\Theta)$, the MLE estimate is

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{(\mathbf{i},\mathbf{j})\in\mathcal{P}} \|r_{\mathbf{i},\mathbf{j}} - p_{\mathbf{i},\mathbf{j}}(\Theta)\|^2 + \text{reg.} \qquad (1)$$

- $r_{\mathbf{i},\mathbf{j}}$ is the empirical average response
- Θ is the model parameter

Let $p_{\mathbf{i}k}$ being the probability that pixel $\mathbf{i}$ belongs to segment k, then

$$p_{\mathbf{i},\mathbf{j}}(\Theta) = \langle p_{\mathbf{i}}, p_{\mathbf{j}} \rangle = p_{\mathbf{i}1}p_{\mathbf{j}1} + \cdots + p_{\mathbf{i}K}p_{\mathbf{j}K}.$$

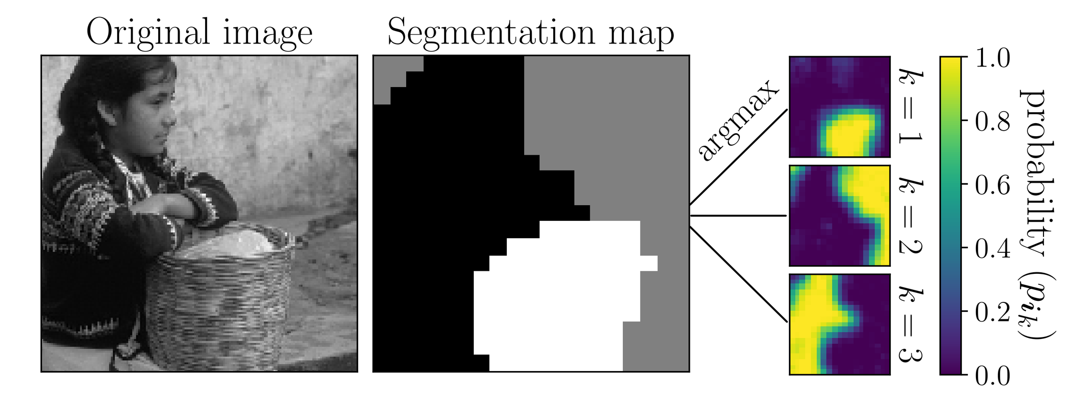
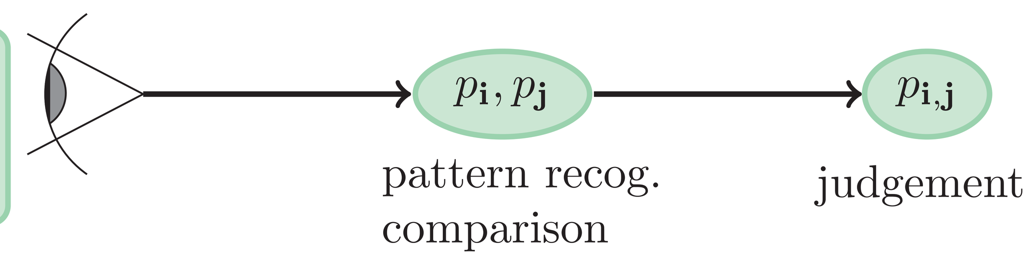


Figure 2: Segmentation map of a natural image and probability of assignment to each segment, obtained with our protocol

4. Models

(i) a non-parametric model $[\Theta = ((p_{\mathbf{i}})_{\mathbf{i}}, \alpha)]$

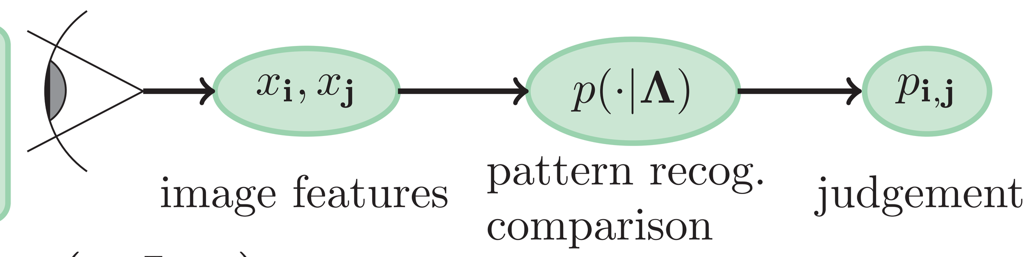
$$p_{\mathbf{i},\mathbf{j}}(\Theta) = \alpha + (1-2\alpha)\langle p_{\mathbf{i}}, p_{\mathbf{j}} \rangle \qquad \text{(NP)}$$



- assumes the existence of underlying probability maps
- $p_{\mathbf{i}} = (p_{\mathbf{i}1}, \dots, p_{\mathbf{i}K})$ with $p_{\mathbf{i}k}$ being the probability that pixel $\mathbf{i}$ belong to segment k
- α: lapse rate

(ii) generative model (optimal observer) $[\Theta = (\Lambda, \alpha)]$

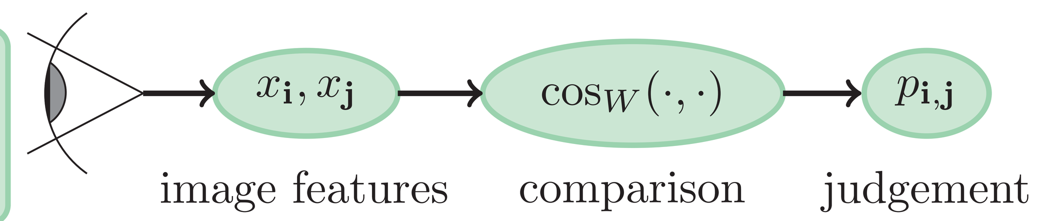
$$p_{\mathbf{i},\mathbf{j}}(\Theta) = \alpha + (1-2\alpha)\langle p(x_{\mathbf{i}}|\mathbf{\Lambda}), p(x_{\mathbf{j}}|\mathbf{\Lambda}) \rangle \qquad \text{(GM)}$$



- assumes the probability maps are obtained via probabilistic inference
- $p_k(x|\mathbf{\Lambda}) \propto \exp\left(-\frac{x^T\Lambda_k x}{2}\right)$ w/ $\Lambda_k = (\Sigma_k + \sigma_0 \mathrm{I})^{-1}$
- Σ_k, σ_0: feature covariance and internal noise

(iii) feedforward discrimininative model $[\Theta = (W, \mu, \sigma, \alpha)]$

$$p_{\mathbf{i},\mathbf{j}}(\Theta) = \alpha + (1-2\alpha)S_{\sigma,\mu}(\cos_W(x_{\mathbf{i}}, x_{\mathbf{j}})) \qquad \text{(FD)}$$



- assumes that local features are directly compared
- $S_{\sigma,\mu}(u) = \left(1 + \exp\left(-\frac{1}{\sigma}\left(\log\left(\frac{u}{1-u}\right) - \mu\right)\right)\right)^{-1}$
- μ, σ: subjective eq. and inverse sensitivity

5. Segmentation Results (4 participants)

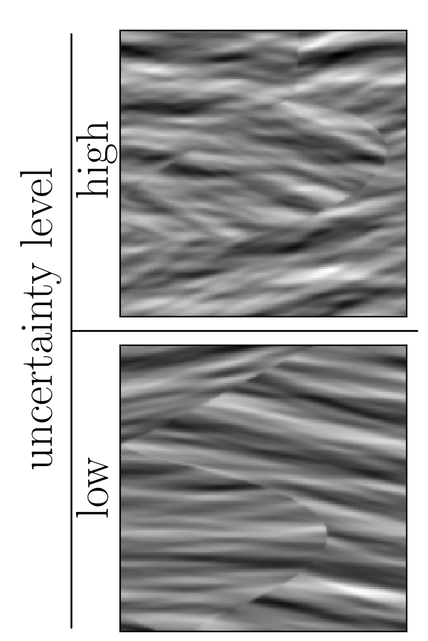


Figure 3: High and low uncertainty stimuli.

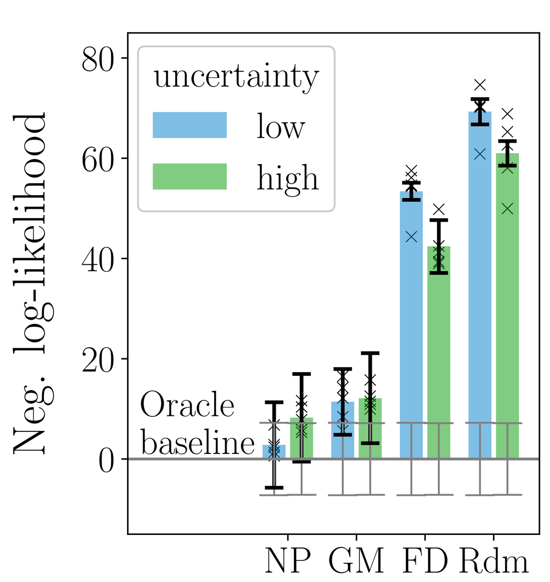


Figure 4: Fit quality (cross-val. negative log-lkl, lower is better). Rdm: chance level.

- manipulating the orientation and spatial frequency distributions of the textured segments changes the segmentation uncertainty – Figure 3
- the probabilistic inference model (GM) explains the data better than the feature discrimination model (FD) – Figure 4
- variability of human segmentation correlates with image uncertainty – Figure 5
- GM captures the variability that is intrinsic to image uncertainty, differences with NP account for other factors such as measurement noise and inter-participants variability – Figure 5
- variability is concentrated around edges, this effect is stronger for low uncertainty stimuli (blue) where edges are more spatially localized – Figure 6

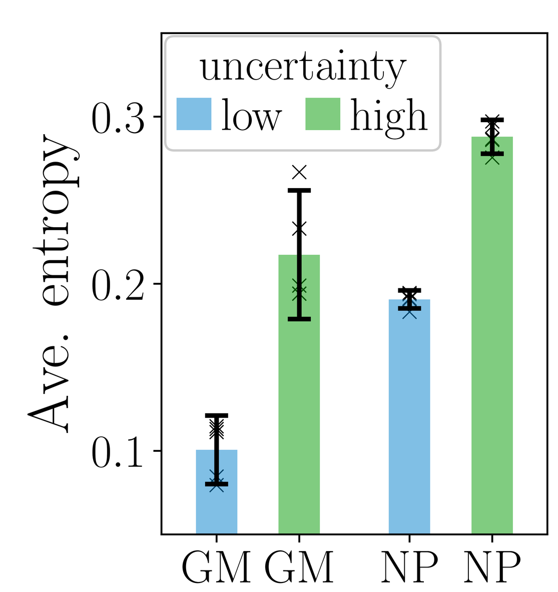


Figure 5: Average entropy.

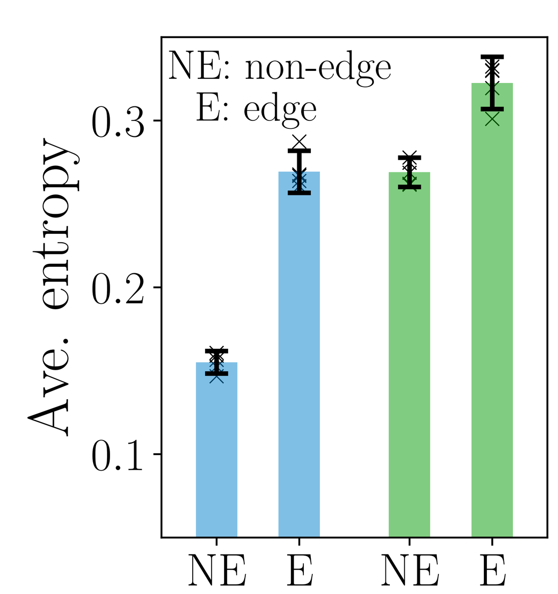


Figure 6: Average entropy for edge and non-edge areas obtained with NP model. Error bars: 99.7% conf. interval.

6. Manipulating Edge Uncertainty

- Edges correspond to jumps of local feature distributions

These jumps are harder to perceive when:

- the distributions are largely overlapping (large bandwidth, top of Figure 7)
- the 1st distribution changes continuously over space towards the 2nd one (edge width, bottom of Figure 7 and Figure 8)

We measured the sensitivity to detect the presence of a vertical edge in an image in a 2AFC (2 images presented side by side). We varied both the edge width and the bandwidth. A trial example is given in Figure 9.

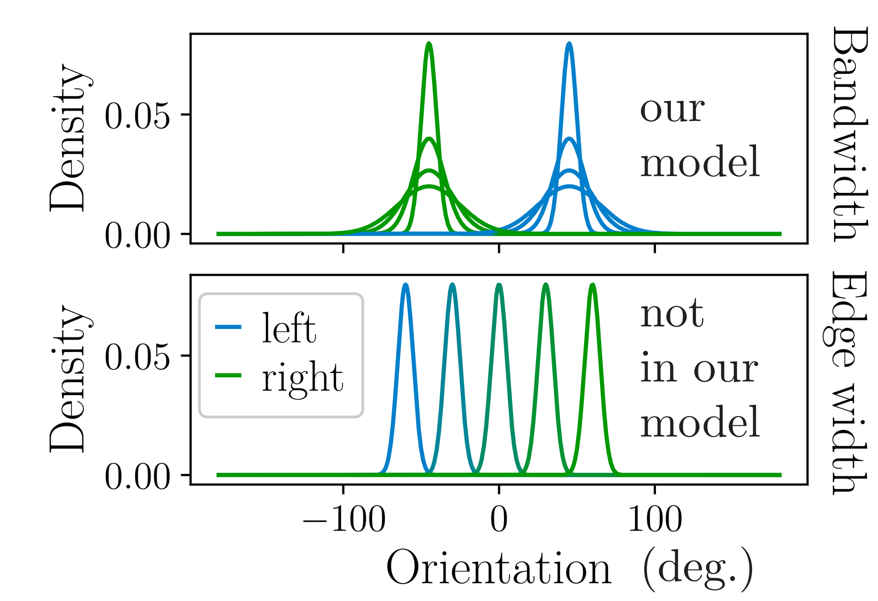


Figure 7: Local distributions of features around an edge

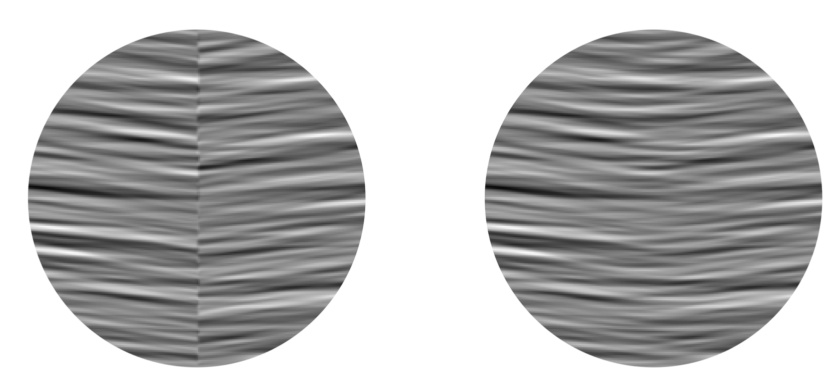


Figure 8: Example of narrow and wide edges

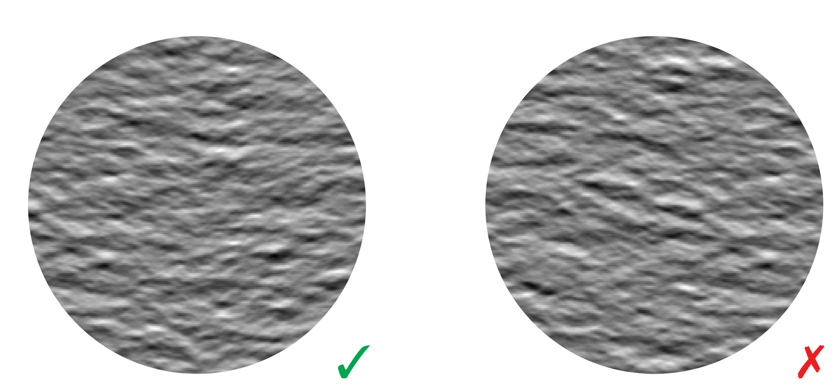


Figure 9: Trial example

Results (pilot, 1 participant):

- no global trend in threshold
- approximately constant sensitivity

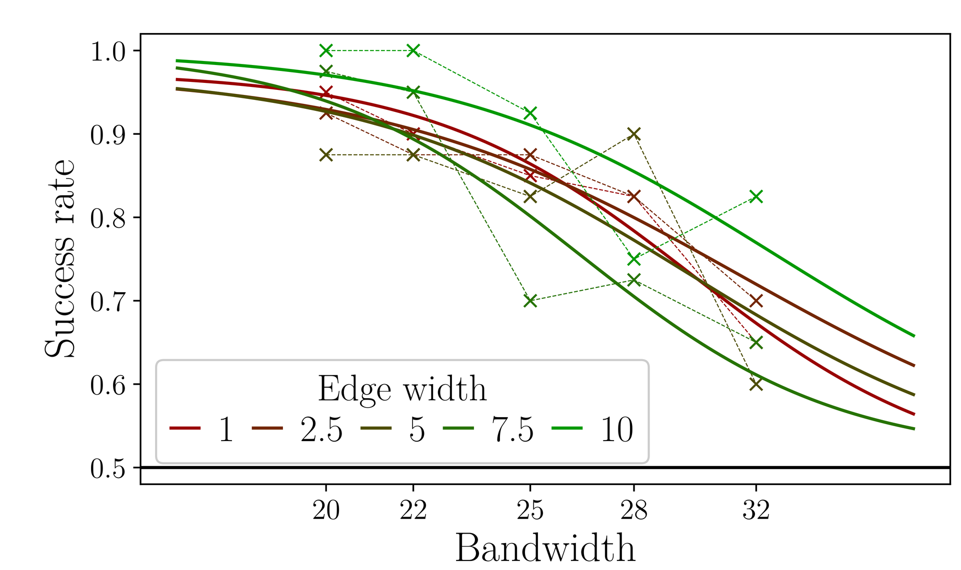


Figure 10: Psychometric function (40 samples)

7. Summary

- we propose a new protocol to study natural image segmentation

Using artificial composite textures, we have found that:

- the generative model accounts better for the data than the discriminative model
- human variability correlates with image uncertainty
- variability is localized around edges

8. Funding

This collaborative work is jointly funded by NIH (NIH/CRCNS EY031166) in USA and ANR in France.

9. References

[1] J. Wagemans et al. In: *Psychological bulletin* 138.6 (2012), p. 1172.

[2] P. R. Roelfsema. In: *Annu. Rev. Neurosci.* 29 (2006), pp. 203–227.

[3] M. S. Landy and J. R. Bergen. In: *Vision research* 31.4 (1991), pp. 679–691.

[4] J. Freeman and E. P. Simoncelli. In: *Nature neuroscience* 14.9 (2011), p. 1195.